

IN THE CLAIMS

Please amend the claims as follows:

Claims 1-18 (Canceled).

Claim 19 (New): A homology analysis system for analyzing whether an analysis target data group is similar to a first data group or a second data group, the first data group having n_A data and the second data group having n_B data, wherein the first and second data groups is different from the analysis target data group, comprising:

amino acid sequence storage means for storing first data groups and second data groups, the first data groups and second data groups being different from the analysis target data group consisting of a gene databank of eukaryotes and the analysis target data group consisting of a gene databank of prokaryotes;

first homology value calculation means for selecting a gene data of the analysis target to compare with each gene data of the gene databank of prokaryotes stored in the amino acid sequence storage means, setting n thresholds E_i ($i=1, 2, \dots, n$) each indicating a determination criterion for the presence/absence of a homology between gene data, wherein the first homology value calculation means determines that a homology is present when a value of a homology between data included in the analysis target data group and the first data group respectively is no less than the thresholds E_i , and calculates, as a first homologous value x_{ij} ($j=1, 2, \dots, n_A$), the number of items of data in which homology is present to calculate a mean value \bar{x}_i of the calculated first homology values x_i with respect to the threshold E_i ;

second homology value calculation means for selecting a gene data of the analysis target to compare with each gene data of the gene databank of prokaryotes stored in the amino acid sequence storage means, setting n thresholds E_i ($i=1, 2, \dots, n$) each indicating a determination criterion for the presence/absence of a homology between gene data, wherein

the second homology value calculation means determines that a homology is present when a value of a homology between data included in the analysis target data group and the second data group respectively is no less than the thresholds E_i , and calculates, as a second homologous value y_{ij} ($j=1, 2, \dots, n_A$), the number of items of data in which homology is present to calculate a mean value $y_{i_}$ of the calculated second homology value y_i with respect to the threshold E_i ;

homology value storage means for storing a homology value hit number and the mean values $x_{i_}$ and $y_{i_}$; and

homology determination means for calculating a homology determination value $Z_i^{(1)}$ indicating similarity to one of the first data group and the second data group according to

$$Z_i^{(1)} = \frac{\bar{x}_i - \bar{y}_i}{u_i} \cdot \sqrt{\frac{nA \cdot nB}{nA + nB}} \quad (i = 1, 2, \dots, n)$$

when

$$u_i = \sqrt{\frac{1}{nA + nB - 2} \left\{ \sum_{j=1}^{nA} (x_{ij} - \bar{x}_i)^2 + \sum_{k=1}^{nB} (y_{ik} - \bar{y}_i)^2 \right\}}$$

wherein the homology analysis system determines whether the analysis target data group is similar to either one of the first data group or the second data group based on the relationship between the first homologous value x_i and the second homologous value y_i , and the number n of the thresholds, and

wherein the first and second homology value calculation means and the homology determination means function as a processor.

Claim 20 (New): A homology analysis system according to claim 19, wherein the processor further comprises determination result derivation unit determining that the analysis target data group has many data having homologies with the first data group, if the homology

determination value $Z_i^{(1)}$ is larger than $t_\alpha(0, 10)$ wherein the homology determination value $Z_i^{(1)}$ is in accordance with a t-distribution and α is a degree of freedom.

Claim 21 (New): A homology analysis system according to claim 19, wherein the processor further comprises determination result derivation means determining that the analysis target data group has many data having homologies with the second data group, if the homology determination value $Z_i^{(1)}$ is smaller than $t_\alpha(0, 10)$ wherein the homology determination value $Z_i^{(1)}$ is in accordance with a t-distribution and α is a degree of

$$Z^{(2)} = \frac{\left| \overline{Z^{(1)}} \right| - t_{nA+nB-2}(0,10)}{s / \sqrt{(n-1)}}$$

freedom.

Claim 22 (New): A homology analysis system according to claim 20, wherein the determination result derivation means further comprises homology validity determination means calculating a homology validity determination value $Z^{(2)}$ given by where \underline{s} is a standard deviation of $Z_i^{(1)}$ and $Z_i^{(1)}_{-}$ is a mean value of $Z_i^{(1)}$, and determining that the homology determination value $Z_i^{(1)}$ is an invalid value, if the homology validity determination value $Z^{(2)}$ is less than a predetermined value $t_{n-1}(0, 10)$.

Claim 23 (New): A homology analysis system according to claim 21, wherein the determination result derivation means further comprises homology validity determination means calculating a homology validity determination value $Z^{(2)}$ given by where \underline{s} is a standard deviation of $Z_i^{(1)}$ and $Z_i^{(1)}_{-}$ is a mean value of $Z_i^{(1)}$,

$$Z^{(2)} = \frac{\left| \overline{Z^{(1)}} \right| - t_{nA+nB-2}(0,10)}{s / \sqrt{(n-1)}}$$

and determining that the homology determination value $Z_i^{(1)}$ is an invalid value, if the homology validity determination value $Z^{(2)}$ is less than a predetermined value $t_{n-1}(0, 10)$.

Claim 24 (New): A homology analysis system according to claim 20, wherein the degree of freedom α is $n_A + n_B - 2$.

Claim 25 (New): A homology analysis system according to claim 21, wherein the degree of freedom α is $n_A + n_B - 2$.

Claim 26 (New): A homology analysis system according to claim 22, wherein the degree of freedom α is $n_A + n_B - 2$.

Claim 27 (New): A homology analysis system according to claim 23, wherein the degree of freedom α is $n_A + n_B - 2$.

Claim 28 (New): A homology analysis system according to claim 19, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 29 (New): A homology analysis system according to claim 20, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 30 (New): A homology analysis system according to claim 21, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 31 (New): A homology analysis system according to claim 22, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 32 (New): A homology analysis system according to claim 23, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 33 (New): A homology analysis system according to claim 24, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 34 (New): A homology analysis system according to claim 25, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 35 (New): A homology analysis system according to claim 26, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 36 (New): A homology analysis system according to claim 27, wherein the first and second homology value calculation means calculates the homology values x_i and y_i by a BLAST method.

Claim 37 (New): A homology analysis system according to claim 19, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 38 (New): A homology analysis system according to claim 20, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 39 (New): A homology analysis system according to claim 21, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 40 (New): A homology analysis system according to claim 22, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 41 (New): A homology analysis system according to claim 23, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 42 (New): A homology analysis system according to claim 24, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 43 (New): A homology analysis system according to claim 25, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 44 (New): A homology analysis system according to claim 26, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 45 (New): A homology analysis system according to claim 27, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 46 (New): A homology analysis system according to claim 28, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 47 (New): A homology analysis system according to claim 29, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 48 (New): A homology analysis system according to claim 30, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 49 (New): A homology analysis system according to claim 31, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 50 (New): A homology analysis system according to claim 32, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 51 (New): A homology analysis system according to claim 33, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 52 (New): A homology analysis system according to claim 34, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 53 (New): A homology analysis system according to claim 35, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.

Claim 54 (New): A homology analysis system according to claim 36, wherein the analysis target data group, the first data group, and the second data group are data representing gene sequences.